A MATHEMATICAL THEORY OF EVIDENCE FOR G.L.S. SHACKLE*

Guido Fioretti

Università di Firenze
and
ICER, Torino

February 2001

Abstract

Evidence Theory is a branch of mathematics that concerns the combination of empirical evidence in an individual's mind in order to construct a coherent picture of reality. Designed to deal with unexpected empirical evidence suggesting new possibilities, evidence theory has a lot in common with Shackle's idea of decision-making as a creative act. This essay investigates this connection in detail, pointing to the usefulness of evidence theory to formalise and extend Shackle's decision theory.

In order to ease a proper framing of the issues involved, evidence theory is not only compared with Shackle's ideas but also with additive and sub-additive probability theories. Furthermore, the presentation of evidence theory does not refer to the original version only, but takes account of its most recent developments, too.

* I wish to thank Victoria Chick, Jean Czerlinski, Marco Dardi, Vittorio Emanuele Ferrante, Thorbjørn Knudsen, Klaus Nielsen, Bart Nooteboom and Germano Resconi for helpful comments and remarks.

1. Unpredictable Hypotheses

Throughout the 1950s and 1960s G.L.S. Shackle was the protagonist of an intense debate about the possibility of adopting a decision theory that, opposing the one suggested by Savage (1954) on the basis of a subjective interpretation of probability (De Finetti, 1931; Ramsey, 1931), would account for creativity and inventiveness in decision-making:

> (...) we think of uncertainty as more than the existence in the decision-maker's mind of plural and rival (mutually exclusive) hypotheses amongst which he has insufficient epistemic grounds of choice. Decision, as we mean the word, is creative and is able to be so through the freedom which uncertainty gives for the creation of *unpredictable hypotheses*. Decision is not choice amongst the delimited and prescribed moves in a game with fixed rules and a known list of possible outcomes of any move or sequence of moves. There is no assurance that any one can in advance say what set of hypotheses a decision maker will entertain concerning any specified act available to him. Decision is thought and not merely determinate response.

(Shackle, 1961, p. 6)

Clearly, the formation of a set of possibilities in the mind of a decision-maker is the crucial step of creative decision-making. Shackle addressed this issue in terms of "degrees of possibility" (Shackle, 1961, p. 12), or equivalently, in order to lay stress on the emergence of novelties in empirical reality, in terms of a "degree of surprise" (Shackle, 1961, pp. 74-76).

More precisely, since this surprise refers to imagined events, Shackle preferred to speak of a "potential surprise":

> It is the degree of surprise to which we expose ourselves, when we examine an imagined happening as to its possibility, in general or in the prevailing circumstances, and assess the obstacles, tensions and difficulties which arise in our minds when we try to imagine it occurring, that provides the indicator of degree of possibility. This is the surprise we *should* feel, if the given thing *did* happen; it is *potential* surprise.

(Shackle, 1961, p. 68)

Shackle stressed that, unlike probability, potential surprise is not a "distributional" uncertainty variable. By a "distributional uncertainty variable" he meant one which assigns a given mass to a given set of possibilities, leaving no room for unpredictable events (Shackle, 1961, pp. 47-53); on the contrary, degrees of potential surprise can be assigned to any number of hypotheses without requiring them to sum up to unity.

Upon the concept of potential surprise Shackle developed a whole decision theory. Unfortunately, lack of a suitable formalisation pushed Shackle's theory to the background, whereas increasing acceptance of subjective probability theories favoured expected utility maximisation as an explanation of decision-making.

The reason for writing this essay is that, after the debate on Shackle's ideas had settled down, a formalisation and generalisation of what is basically the same concept was made available by a mathematician, Glenn Shafer (1976), and that this formalisation remained unnoticed among economists hitherto. The converse is not true, since the work of the British economist received considerable attention by Shafer (Shafer, 1976, p. 223); however, possibly because deeply involved in the refinement of mathematical details, Shafer never bothered to popularise his theory among economists.

Since 1976 evidence theory is gaining increasing acceptance in the field of artificial intelligence, particularly in the design of expert systems (Shafer, 1990, 1992). On the contrary, its popularity among probabilists remains low, since evidence theory requires epistemological assumptions that are at odds with those underlying probability theories.

More than on technical details, where evidence theory is actually very similar to the theories of sub-additive probabilities, the difference with probability theories lies in the fact that evidence theory understands possibilities as conceived in the mind of decision-makers, rather than objectively given in a "possibility set". Evidence theory can deal with the continuous formation of "unpredictable hypotheses" in the decision-makers' minds, suggested by the empirical evidence they receive. Non-additive probability theories can yield similar numerical results as the ones evidence theory arrives at, but only under the unrealistic assumption of choosing *ex ante* a particular possibility set.

The aim of this essay is to explain the basic ideas of evidence theory without sticking on technicalities that can be found elsewhere, to frame these ideas with respect to probability theory and, above all, to establish a link between Shafer's "evidence theory" and the foundations of Shackle's decision theory. In particular, section 2 is devoted to expound the theory of sub-additive probability and to explain why it is not able to capture uncertainty as Shackle understood it, while section 3 utilises this background to illustrate the basic concepts of evidence theory. Subsequently, section 4 compares Shafer's and Shackle's uncertainty theories, while section 5 compares their respective decision theories.

2. Sub-Additive Probabilities

Let A and B denote two urns. Let us suppose that a person who extracts one ball from each urn knows that urn A contains an equal number of white and black balls, while all she knows about urn B is that it contains white and black balls. According to Laplace's principle of insufficient reason, this person must assume in both cases equal probabilities $1/2$ to extract a white or a black ball. Yet it is intuitively evident that the uncertainty relative to urn A is not the same as the uncertainty relative to urn B.

Invented by von Kries as a criticism of the idea of attaching equal prior probabilities to all conceivable events, this famous paradox was resurrected by Ellsberg after almost a century as an example of a situation where subjective probability theory leads to absurd results (von Kries, 1886; Ellsberg, 1961). The situation it describes conveys a feeling that there exists a kind of uncertainty that probability is not able to capture, a kind of uncertainty that Ellsberg called "ambiguity".

Ellsberg's paradox is not so destructive for subjective probability theory, however. Its appeal derives from the fact that it compares two extreme cases: in the first case the person who extracts the ball knows the exact proportion of white and black balls that are in the urn, a knowledge that is equivalent to saying that she measured probabilities on a sample of infinite dimension; in the second case the person who extracts the ball has no idea of the ratio of white to black balls in the urn, which amounts to say that she has no sample at all to measure probabilities or, equivalently, that the sample has zero dimension.

The real issue at stake may become more evident if we look at intermediate cases, where the dimension of the sample is neither infinite nor zero. The following example is due to Dempster:

> In order to illustrate the idea simply, consider a map showing regions of land and water. Suppose that $0{\cdot}80$ of the area of the map is visible and that the visible area divides in the proportions $0{\cdot}30$ to $0{\cdot}70$ of water area to land area. What is the probability that a point drawn at random from the *whole* map falls in the region of water ? Since the visible water area is $0{\cdot}24$ of the total area of the map, while the unobserved $0{\cdot}20$ of the total area could be water or land, it can be asserted only that the desired probability lies between $0{\cdot}24$ and $0{\cdot}44$.

(Dempster, 1968)

This example is meant to illustrate the idea of expressing uncertainty by means of probability intervals, instead of by means of single probability values. Basically, the idea is that the width of a probability interval can be used to represent the dimension of the sample: for instance, the appeal of Ellsberg's paradox derives from the fact that it compares the extreme case of a zero-dimensional sample, where the probability interval is the whole $[0,1]$, with the opposite extreme of a sample of infinite dimension, where the probability interval shrinks down to a single point at 0.5.

The idea of expressing uncertainty by means of probability intervals dates back to Koopman (1940a, 1940b, 1941), who denoted the extremes of a probability interval $\left[p_*, p^*\right]$ as 'lower probability' p_* and 'upper probability' p^*. More recently, lower and upper probabilities have been derived from betting rates, in a setting that is more akin to that of subjective probability theory (Smith, 1961, 1965; Williams, 1976).

Note that the mental habit of subjective probability theory could suggest to attach a subjective probability distribution to the values of a probability interval, and to reduce a probability interval to a single probability value by taking the expected value of this distribution. This is not the case, however, since this probability distribution can produce an expected probability interval instead of a single value; more in general, a probability interval $\left[p_*, p^*\right]$ should be understood as the first step of an infinite hierarchy of increasingly fuzzy probability intervals and probability distributions (Good, 1952, 1962).

The basic relationships of ordinary probability theory become:

$$p_*(\Theta) = p^*(\Theta) = 1 \tag{1}$$

$$p_*(\varnothing) = p^*(\varnothing) = 0 \tag{2}$$

$$p_*(A) + p^*(\overline{A}) = 1 \tag{3}$$

where Θ is the set of all possible events, $\varnothing$ is the void set, $A \in \Theta$ and $\overline{A}$ is the complement of A. Lower and upper probabilities can be combined according to rules that, in the particular case of $p_* = p^* = p$, reduce to Bayesian conditioning (Dempster, 1966, 1967, 1968).

Relation (3) suggests that lower probabilities do not sum up to unity. In fact, since $p_*(A) \le p^*(A)$, it is immediate to see that

$$\sum_i p_*(A_i) \leq 1 \qquad (4)$$

In fact, lower probabilities do not sum up to unity when they are measured upon too small a sample. Thus, they are able to condense the usual concept of probabilistic uncertainty and Ellsberg's "ambiguity" in a single magnitude.

Lower probabilities, also called "non-additive" or "sub-additive" probabilities, are a more general and flexible tool to express probabilistic uncertainty than additive probabilities are (Walley 1991). Furthermore, by means of capacity theory (Choquet 1953-54) it has been possible to extend the concept of subjective expected utility to sub-additive probability distributions (Gilboa, 1987; Schmeidler, 1989; Sarin and Wakker, 1992; Jaffray and Philippe, 1997). And even the "decision weights" of prospect theory (Kahneman and Tversky, 1979), a generalisation of expected utility that can account for many of its empirical failures, can be understood as sub-additive probabilities (Tversky and Kahneman, 1992; Sarin and Wakker, 1994).

Most important, since sub-additive probabilities do not necessarily sum up to unity, they are non-distributional variables. However, this does not mean that non-additive probabilities are able to express the "creation of unpredictable hypotheses" Shackle spoke of.

In order to do so, we must abandon the very idea that it is possible to list all events in a "possibility set", including the future ones. As we shall see in the next section, evidence theory marks a fundamental turning point in the modellisation of uncertain reasoning, since it substitutes the "possibility set" with a "frame of discernment" that reflects the decision-maker's cognitive capabilities.

3. The Frame of Discernment

Any theory of probability takes some possibility set as given. The events defined on it might be nested in one another or they might partially overlap, complex hierarchies of events can be envisaged, but in any case, the possible subdivisions of the possibility set into subsets inevitably come to an end at a certain point. Let us call 'elementary' the events that are identified by the finest subdivision of the possibility set.

The very existence of elementary events, by whose combination any other event can be constructed, implies that Shackle's "unpredictable hypotheses" are out of the scope of any probability theory. Nonetheless, a great probabilist like De Finetti did acknowledge that any identification of elementary events, any exclusion of "unpredictable hypotheses" from the outset, is necessarily arbitrary and provisional:

> We interpreted these elements of the “last subdivision” as “points”, but any idea which does not take account of the relative, arbitrary and provisional character of this stop in subdividing, which thinks them as “indivisible” or as “less subdivisible”, or however different from all other events, is groundless and misleading.
>
> (De Finetti, 1970, p. 44)

Evidence theory does not assume that a possibility set is objectively given, but only that individuals, at a certain point in space and time, are able to discern certain events from empirical evidence. In order to stress its cognitive nature, Shafer even preferred to call the set of the possibilities envisaged by an individual with another name:

> It should not be thought that the "possibilities" that comprise [a set] Θ will be determined and meaningful independently of our knowledge. Quite to the contrary: Θ will acquire its meaning from what we know or think we know; the distinctions that it embodies will be embedded within the matrix of our language and its associated conceptual structures and will depend on those structures for whatever accuracy and meaningfulness they possess. In order to emphasize this epistemic nature of the set of possibilities Θ, I will call it the *frame of discernment*. When a proposition corresponds to a subset of a frame of discernment, I will say that the frame *discerns* that proposition.

(Shafer, 1976, p. 36)

A frame of discernment does not know anything like "elementary events". On the contrary, the frame of discernment changes any time an individual thinks "unpredictable hypotheses" according to the problem to be solved:

> In the past, many students of probable reasoning have sought to establish a fixed framework for their speculations by postulating the existence of an ultimately detailed set of "possible states of nature" - a frame of discernment so fine that it encompasses all possible distinctions and admits of no further refinement. (...) we must reject the postulation of such an ultimate refinement.
>
> This rejection is compelled by the purely epistemic nature of the role played by the frame of discernment. In our theory, one's frame of discernment is not a set of "states of nature" that are objectively possible in the sense that they are allowed by some physical law. Nor is it a set of situations that one might recognize as distinct possibilities if one knew more than one does. It is a set of possibilities that one does recognize on the basis of knowledge that one does have - or at least on the basis of distinctions that one actually draws and assumptions that one actually makes. It cannot embody concepts and distinctions that one has never heard of. Yet, as we see after a moment's reflection, each "possible state of nature" in an ultimate refinement would have to be a complete account of the history of the universe, the very writing of which would require knowledge surpassing the collective experience of mankind.
>
> In practice, we always begin with a relatively coarse frame of discernment and then refine it as thought and evidence require, often producing distinctions that were outside the scope of our attention when we began. Hence a realistic theory of evidence will deal with frames that do not even encompass all the knowledge we do have and will explicitly allow for their refinement.

(Shafer, 1976, p. 119)

Evidence theory does not deal with the cognitive processes whereby individuals change their frames of discernment, which amounts to say that it does not deal with the way "unpredictable hypotheses" arise. Rather, evidence theory assumes that, for a certain individual and at a certain point in time, a provisional frame of discernment is given, and it investigates which beliefs this individual forms out of the empirical evidence he receives.

Empirical evidence is assumed to be available as sets of numbers $\{m(A_1), m(A_2), m(A_3)\dots\}$ which represent the amounts of evidence that support hypotheses $\{A_1, A_2, A_3 \dots\}$. Eventually, a number $m(\Theta)$ committed to frame of discernment Θ means that the evidence committed to $\{A_1, A_2, A_3 \dots\}$ is inconclusive. Each set of numbers $\{m_1, m_2, m_3, \dots\}$ is called a *body of evidence*: one body of evidence is relative to $\{A_1, A_2, A_3 \dots \Theta\}$, another to $\{B_1, B_2, B_3 \dots \Theta\}$, and so on.

In order for evidence theory to entail probability theory as a particular case it is necessary to require that, for any given body of evidence, numbers $\{m\}$ sum up to unity:

$\sum_i m(A_i) + m(\Theta) = 1$, $\sum_i m(B_i) + m(\Theta) = 1$, and so on.[(1)] Note that the presence of $m(\Theta)$ eventually allows the rest of the evidence not to sum up to one: in this way we can take account of imperfect information, just like sub-additive probabilities do. Note also that the normalisation to unity is not in contrast with Shackle's call for a non-distributional uncertainty variable, since new hypotheses would be supported by a new body of evidence, and the normalisation of the corresponding $\{m\}$ would not affect the previous evidence bodies.

Unlike possibility sets of probability theory, a frame of discernment is not an algebra. This means that its elements cannot be combined according to rules (e.g. union or intersection of subsets) that are independent of the individual's frame of discernment itself. In other words, combination of empirical evidence must be carried out according to the structure of the frame of discernment, yielding beliefs exclusively on the issues an individual can conceive and not on any possible combination of "elementary" events.

This is a point of the highest importance, actually. It means that a frame of discernment cannot support an algorithm that combines possibilities so as to produce more structured concepts. Rather, a frame of discernment entails the possibilities that an individual holistically distinguishes from one another, without assuming that his mind combines these possibilities according to given rules.

The objects a mind focuses its attention on, are the *focal elements* of the corresponding frame of discernment: they are indivisible entities, that cannot be understood as a combination of more fundamental components. For instance, a firm is not only a set of machineries plus some amount of labour force: its habits and its culture make it an organisation, an object that is different from the sum of its parts. Or to cite a similar example in a completely different field: since the focal elements of a frame of discernment are indivisible quanta, evidence theory is the proper framework for the kind of uncertainty that arises in quantum theory (Resconi, Klir and Pessa, 1999).

Evidence theory represents individuals' beliefs by means of *belief functions*. An individual is not assumed to have beliefs on any proposition P, but only on those which his frame of discernment can identify, and the evidence that supports P is only that one which bears on the subsets of P. Thus, given a frame of discernment Θ and a body of empirical evidence $\{m(A_1), m(A_2), m(A_3) \ldots\}$, the belief committed to $P \in \Theta$ is:

$$Bel(P) = \sum_{A_i \subset P} m(A_i) \qquad (5)$$

Belief functions, like sub-additive probabilities, are non-distributional variables that do not sum up to unity. It is conventionally assumed that $Bel(\Theta) = 1$, but this belief cannot be distributed to the subsets of Θ.

Given several belief functions over the same frame of discernment, but based on distinct bodies of evidence, Dempster-Shafer rule enables us to compute a new belief function based on the combined evidence (Dempster, 1966, 1967, 1968; Shafer, 1976). Let $m' = \{m(A_1), m(A_2), m(A_3) \ldots\}$ and $m'' = \{m(B_1), m(B_2), m(B_3) \ldots\}$ denote two distinct bodies of empirical evidence, that support belief functions Bel' and Bel'', respectively. The following combined evidence supports the combined belief function $Bel = Bel' \oplus Bel''$ on P:

$$m(P) = \frac{\sum_{A_i \cap B_j = P} m'(A_i)\ m''(B_j)}{1 - \sum_{A_i \cap B_j = \varnothing} m'(A_i)\ m''(B_j)} \qquad (6)$$

where $\varnothing$ represents the void set.

Dempster-Shafer combination rule allows evidence theory to take account of Shackle's "unpredictable hypotheses": in fact, the new propositions that arise in the mind of the decision-maker must be supported by a new body of evidence, and this new body of evidence can be combined with the old ones by means of (6). Notice, however, that Bayes's theorem can be shown to be a particular case of Dempster-Shafer rule: thus, the ability of evidence theory to take account of "unpredictable hypotheses" is due to its epistemological premises, rather than to its mathematics.

Two examples follow, in order to explain the above considerations. The first example illustrates the functioning of evidence theory in a very simple setting, where evidence theory and probability theory end up to coincide. The second example, on the contrary, concerns a situation where belief functions yield different numerical results from sub-additive probabilities.

Example 1: Tossing a fair coin (adapted from (Shafer, 1976, p. 197)[(2)])

Let the frame of discernment be $\Theta = \{H, T\}$, where focal elements H and T mean "heads" and "tails", respectively. This frame of discernment is constituted by two singletons; as we shall see, in this case belief functions can be reduced to probability functions.

The information obtained by tossing the coin is the number of times it turned out "head" and the number of times it turned out "tail", denoted by n_H and n_T, respectively. If we assume not to know the number of times the coin has been tossed, these numbers are two distinct bodies of evidence.

Let us transform these evidence bodies into $m' = 1 - e^{-n_H}$ and $m'' = 1 - e^{-n_T}$, respectively: evidence is zero $(m = 0)$ if the coin is never tossed $(n = 0)$, it is one $(m = 1)$ if the coin is tossed infinite times $(n \to \infty)$. Evidence m' bears on $H \subset \Theta$; if $n_H < \infty$, and consequently $m' < 1$, the lack of evidence $1 - m'$ bears on Θ as a whole. Obviously, the same holds for m'', T and n_T.

Applying Dempster-Shafer rule (6) we obtain that the evidence in favour of obtaining heads or tails is, respectively:

$$m(H) = \frac{m'(1 - m'')}{1 - m'm''}$$

$$m(T) = \frac{m''(1 - m')}{1 - m'm''}$$

If $n_H = n_T = n$ we can write that, in the limit, infinite evidence yields $\lim_{n\to\infty} m(H) = \lim_{n\to\infty} m(T) = 1/2$, while for $n < \infty$ $m(H) = m(T) < 1/2$. Normalisation $m(H) + m(T) + m(\Theta) = 1$ holds in any case, but only if there is an infinite amount of evidence $m(H)$ and $m(T)$ sum up to one and $m(\Theta) = 0$.

Since frame of discernment $\Theta = \{H, T\}$ is constituted by singletons, formula (5) yields that the beliefs to get a head or a tail by the next toss are $Bel(H) = m(H)$ and $Bel(T) = m(T)$, respectively. Thus, both beliefs increase from zero to $1/2$ with increasing empirical evidence, just like sub-additive probabilities would do.

By endowing the frame of discernment $\Theta = \{H,T\}$ with the usual set operators it is possible to consider beliefs for the combined event $H \cup T$: by doing so Θ becomes an algebra, which we call "possibility set". Thus, belief functions $Bel(H)$ and $Bel(T)$ satisfy $Bel(H \cup T) = Bel(H) + Bel(T)$ and become probability functions $p(H)$ and $p(T)$, respectively. If we toss the fair coin infinite times, or if we know from the outset that it is a fair one, these probabilities are such that $p(H) + p(T) = 1$. Otherwise, $p(H) + p(T) < 1$.

If the frame of discernment boils down to a possibility set, Dempster-Shafer rule boils down to Bayes's theorem. In order to see this in our example, let us understand $p(H)$ as the probability of observing head conditioned on the probability of observing either head or tail, and let us denote this event by #: $p(H) \equiv p(H|\#)$. According to Bayes's theorem, $p(H|\#) = p(H \cap \#) / p(\#)$. But $p(H \cap \#)$ is given by the Cartesian product of the evidence of observing a head and the evidence of not observing a tail, which is $m'(1 - m'')$. In its turn, $p(\#)$ is one minus the eventual contradictory evidence of observing a head and a tail at the same time, that is $1 - m'm''$.

Since evidence theory explicitly represents an individual's cognitive capabilities by means of his frame of discernment, it enjoys a more general status than frequentist and subjectivist probability theories. If the frame of discernment is so simple that it can be reduced to a possibility set - as it is in this example - we face the following option: either attribute the possibility set to the objective reality, in which case we give a frequentist interpretation to the probabilities above, or attribute the possibility set to the individual, in which case the probabilities above are measures whereby the individual updates his priors.

Unfortunately, this makes subjective probability theory unduly dependent on the structure of the possibility set an individual is supposed to have from the outset, as the following example will show. Evidence theory, on the contrary, allows decision-makers to adapt their frames of discernment to changing circumstances.

Example 2: The burglary in the rooming house (adapted from (Shafer, 1981)[(3)])

Suppose that a burglar is traced to a rooming house, in such a way as to make it highly probable that he is actually one of the roomers. A police detective searches the rooming house and interviews the five roomers, but on this first examination finds nothing that either exonerates or further incriminates any of them.

At this point the detective might formulate a frame Θ which includes a subset A_1 corresponding to the possibility that one of the roomer is the burglar, and - say - he might express the available empirical evidence as $m' = \{m(A_1), m(\Theta)\}$ with $m(A_1) = 0.95$ and $m(\Theta) = 0.05$. The corresponding belief function, on a proposition P that means guilt, is:

$$Bel'(P) = \begin{cases} 0 & \text{if } A_1 \not\subset P \\ 0.95 & \text{if } A_1 \subset P \neq \Theta \\ 1 & \text{if } P = \Theta \end{cases}$$

where the value $Bel(\Theta) = 1$ stems from the very definition of belief function.

Let us now suppose that two of the five roomers produce airtight alibis. The inspector changes his frame of discernment in order to take account of this new evidence. This can be formalised as $m'' = \{m(B_1), m(\Theta)\}$, with $m(B_1) = 1$ and $m(\Theta) = 0$. The corresponding belief function is:

$$Bel''(P) = \begin{cases} 0 & \text{if } B_1 \not\subset P \\ 1 & \text{if } B_1 \subset P \neq \Theta \\ 1 & \text{if } P = \Theta \end{cases}$$

The combination of these two belief functions according to the Dempster-Shafer rule (6) yields:

$$Bel(P) = \begin{cases} 0 & \text{if } A_1 \cap B_1 \not\subset P \\ 0.95 & \text{if } A_1 \cap B_1 \subset P \not\supset B_1 \\ 1 & \text{if } B_1 \subset P \end{cases}$$

meaning that the guilt suspicion is moved to the remaining three roomers.

Let us now analyse this detective story by means of probability theory. Let us represent the initial evidence by means of a sub-additive probability of 0.95 that the burglar is one of the roomers. Let us distribute this probability among the five roomers, so that each roomer has probability 0.19 of being the burglar.

Let us update the probability of being the burglar after two roomers provided alibis. The probability of being a roomer and not to have an alibi is $0.19 \times 3 = 0.57$. The probability not to have an alibi is $1 - 0.19 \times 2 = 0.62$. Thus, Bayes's theorem yields a probability $0.57/0.62 \cong 0.92$ for one of the three remaining roomers to be the burglar, which is close to the value 0.95 yielded by evidence theory.

Notice, though, that this analysis is sensitive to the number of roomers and the proportion with alibis in a way that the analysis using belief functions is not. If four out of five roomers have alibis, then the final probability for the remaining one would be only $0.19/0.24 \cong 0.79$; if there were 20 and 19 were similarly exonerated, then the final probability for the remaining one would be $0.0475/0.0975 \cong 0.49$. And these figures could easily be altered if we claimed that our initial evidence justified unequal prior chances for the roomers.

Evidence theory, on the contrary, does not need to distinguish the five roomers from one another from the outset. Only after two of them provided an alibi it is meaningful for the inspector to distinguish them from the remaining three, and only further evidence could lead the inspector to focus on one of them. Note that the inspector changes his frame of discernment each time empirical evidence allows him to restrict the range of suspects, while probability theory requires the inspector to think all possibilities from the outset.

The last example aimed to stress that belief functions and sub-additive probabilities are not the same thing. The difference is not due to the combination rule, however, since Dempster-Shafer rule applies to sub-additive probabilities, too (Jaffray, 1992; Gilboa and Schmeidler, 1993). Rather, the difference stems from the fact that evidence theory allows the individual to modify his frame of discernment according to the possibilities that he is able to conceive, while probability theory requires the individual to know an exhaustive list of elementary events and to conceive more complex events only by combining the elementary ones according to pre-specified rules.

At a deeper level of understanding we can say that probability theory and evidence theory refer to different conceptual models of the way uncertain reasoning works: while probability theory expresses uncertainty in terms of knowledge about chances that govern the truth, evidence theory expresses uncertainty in terms of knowledge provided by a message whose truth is unknown (Shafer, 1981, 1982). In the first case, the possibility set says where the truth can lie; in the second case, the frame of discernment says how an individual can interpret a message, trying to approximate the truth by eliminating contradictions.

Probability theory tends to understand uncertainty in terms of similarity of the situation at hand with a game of chance; evidence theory, on the contrary, resumes an approach that dates back to the end of the 17th century, which understands uncertainty in terms of similarity of the situation at hand with the evaluation of (possibly contradictory) testimonies in a trial (Shafer, 1986b). Originated by George Hooper in 1699 and improved by Johann Heinrich Lambert in 1764, this trial-based approach was obscured by Laplace in the 19th century, re-discovered in the early 1960s by Per Olof Ekelöf, a Swedish professor of law, and formalised by Arthur Dempster in the late 1960s, who did not realise all the implications of what he was doing, however.

Evidence theory and probability theory refer to alternative canonical examples that suit particular decision problems to a greater or lesser extent (Shafer and Tversky, 1985), and the canonical example evidence theory refers to is surely better suited to deal with surprising events and the need to conceive new possibilities. The next section clarifies the link between Shackle's "potential surprise" and Shafer's "belief functions", highlighting commonalties as well as differences between them.

4. Potential Surprise and Belief Functions

Shackle's "potential surprise" is not a belief committed to a certain hypothesis, but to the "least surprising" among its alternatives:

> We would say that an individual's degree of belief in a hypothesis can be easily and exactly expressed by means of the potential surprise he assigns to the least (potentially) surprising rival hypothesis.

(Shackle, 1961, p. 71)

In fact, Shackle found more natural to measure uncertainty by means of a variable expressing "disbelief" than by means of a variable expressing "belief":

> We must once again insist that when the uncertainty in a person's mind arises from the plurality of the answers which suggest themselves to him for some one question, these answers are rivals mutually excluding each other. To believe in one of these answers is therefore to disbelieve in the others. By contrast it is *not* true that to disbelieve in one answer is to believe in the others. Thus it seems more natural, when we require the notion of uncertainty-variable in order to label various answers with this status or that in relation to 'certainty of rightness' or 'certainty of wrongness' to use a variable expressing *disbelief*. Zero potential surprise expresses *zero disbelief*.

(Shackle, 1961, p. 74)

The key sentence of this last passage is "To believe in one of these answers is therefore to disbelieve in the others": Shackle seems to think that a variable expressing belief would necessarily be a distributional one. However, a few lines later he states that his move was only suggested by the fear that a variable expressing belief could be easily turned into a distributional uncertainty variable:

> To invert a problem is a well-known resource of the mathematician. We too, in adopting as our uncertainty-variable a measure of *disbelief*, are in a way inverting our problem, and we are thereby guaranteeing our solution against any attempt to turn it back into a distributional variable. Formally, it would be open to us to define a measure of possibility directly. A zero value of this measure would stand for impossibility, an absolute maximum value for perfect possibility. But it would then be tempting for those who favour a distributional solution to add together (contrary to reason and to the nature of the problem) the respective degrees of possibility assigned to the members of an exhaustive set of hypotheses, and to treat the resulting total as unity. This total could then, at no greater cost in perversity, be looked upon as 'distributed' over the various rival hypotheses. By contrast, when we represent perfect possibility by a zero value of a measure of disbelief, no such summation can in general be meaningfully carried out, since, for example, the sum of a set of zero values is zero.

(Shackle, 1961, p. 75)

Thus, Shackle's idea of uncertainty can be either conveyed by a variable expressing belief or by a variable expressing disbelief, although Shackle himself had a preference for this second solution.

Shafer identified Shackle's "potential surprise" with $Bel(\overline{\mathrm{P}})$: since $\overline{\mathrm{P}}$ is the set of all subsets of Θ except P, $Bel(\overline{\mathrm{P}})$ measures the disbelief in P and, at the same time, the surprise that would be felt if any $\mathrm{Q} \in \overline{\mathrm{P}}$ would occur (Shafer 1976, p. 225). Several writers in Shackle's vein, on the contrary, turned to non-distributional variables that express belief instead of disbelief: thus, Katzner's "potential confirmation" (Katzner, 1986) or Ford's "degree of credibility" (Ford, 1987) should be understood as $Bel(\mathrm{P})$. But apart from a preference for focusing on belief or disbelief, the important fact is that Shackle's idea of a "potential surprise" can be expressed by means of Shafer's "belief functions".

Furthermore, the slight difference between Shackle and Shafer as far as it concerns "belief" and "disbelief" disappears as soon as both authors advocate the use of two non-distributional variables at the same time. Shackle wrote that an individual's belief should be expressed by the potential surprise of an hypothesis together with the potential surprise of its contradictory:

When we consider together as one whole all the hypotheses which are rivals to a given hypothesis H, they constitute the *contradictory* of H. If, then, we wish to express the individual's *degree of belief* in a hypothesis, we shall think of this degree of belief as consisting in a degree of potential surprise associated with the hypothesis, and in another degree associated with its contradictory.

(Shackle, 1961, p. 73)

Shafer, in his turn, wrote pretty much in the same vein:

One's beliefs about a proposition P are not fully described by one's degree of belief $Bel(\mathrm{P})$, for $Bel(\mathrm{P})$ does not reveal to what extent one doubts P - i.e., to what extent one believes its negation $\overline{\mathrm{P}}$. A fuller description consists of the degree of belief $Bel(\mathrm{P})$ together with the *degree of doubt* $Dou(\mathrm{P}) = Bel(\overline{\mathrm{P}})$. [(4)]

(Shafer, 1976, p. 43)

Thus, it appears that both Shackle and Shafer express uncertainty by means of $Bel(\mathrm{P})$ and $Dou(\mathrm{P})$, the only difference being that Shackle focused on the degree of doubt while Shafer focused on the degree of belief.

The rule employed by Shackle to combine different bodies of empirical evidence seems to be very different from Dempster-Shafer rule (6), at first sight:

The degree of potential surprise associated with any hypothesis will be the least degree amongst all those appropriate to different mutually exclusive sets of hypotheses (each set considered as a whole) whose truth appears to the individual to imply the truth of this hypothesis.

(Shackle, 1961, p. 80)

However, Shafer devotes a whole chapter to the particular case of non-conflicting evidence. In this case belief functions are said to be 'consonant', and their properties are expressed by the following theorem:

Suppose Θ is a finite set. Then a function $Bel: 2^{\Theta} \rightarrow [0, 1]$ is a consonant belief function if and only if $Bel(\Theta) = 1$ and the function $Dou: 2^{\Theta} \rightarrow [0, 1]$ given by $Dou(\mathrm{P}) = Bel(\overline{\mathrm{P}})$ obeys these two rules:

(1) $Dou(\mathrm{P}) = \min_{\theta \in \mathrm{P}} Dou(\{\theta\})$ for all non-empty $\mathrm{P} \subset \Theta$.

(2) There exists $\theta \in \Theta$ with $\mathrm{Dou}(\{\theta\}) = 0$.

(Shafer, 1976, p. 224)

Point (1) is equivalent to Shackle's combination rule quoted above, while point (2) corresponds to the following of Shackle's axioms:

At least one member of an exhaustive set of rival hypotheses must carry zero potential surprise.*

* But it is possible for all the rival hypotheses which are in any degree particularized or specified to carry potential surprise greater than zero, only the residual hypothesis carrying zero potential surprise.

(Shackle, 1961, p. 81)

Thus, Shackle's combination rule appears to be a particular case of Dempster-Shafer rule, one that obtains when empirical evidence is not contradictory. On this subject, Shafer comments the following passage by Shackle:

> To assign greater than zero degrees of potential surprise to both the hypothesis and its contradictory would (...) betray an unresolved mental confusion.
>
> (Shackle, 1961, p. 74 [Shafer, 1976, p. 225])

By no means impressed by the possibility of "unresolved mental confusion", Shafer objects:

> It is easy to share the desire of these scholars to ban the appearance of conflict from our assessment of evidence and our allocation of belief. But in light of what we have learned, the ambition of doing so must be deemed unrealistic. The occurrence of outright conflict in our evidence should and does discomfit us; it prompts us to re-examine both our evidence and the assumptions that underlie our frame of discernment with a view to removing that dissonance. But this effort does not always bear fruit - at least not quickly. And using all the evidence often means using evidence that is embarassingly conflicting.
>
> (Shackle, 1976, p. 225)

It might be interesting to note that, although there have been very few attempts at an empirical validation of Shackle's theory, the limited evidence at our disposal seems to support the usefulness of the concept of "potential surprise" at least in some decision settings (Hey, 1985; Ford and Ghose, 1994a, 1995a, 1995c), but it tends to discard the rest of Shackle's axioms, particularly those concerned with the rules for the combination of evidence (Ford and Ghose, 1994b, 1995b).

Thus, it appears that Shackle actually covered a subsection of evidence theory, the one which deals with non-conflicting evidence. This branch of evidence theory now takes the name of "possibility theory" (Dubois and Prade, 1988), since in this case belief functions can be understood as possibility measures and Dempster-Shafer rule coincides with the rule that is used to combine degrees of membership to fuzzy sets. Evidence theory can be linked to fuzzy sets theory, namely (Klir and Yuan, 1995), and both of them can be seen as deriving from modal logic (Resconi, Klir and St.Clair, 1992; Klir, St.Clair and Harmanec, 1993; Harmanec, Klir and Resconi, 1994; Klir and Harmanec, 1994; Resconi, Klir, Harmanec and St.Clair, 1995).

The connection between evidence theory and the theory of fuzzy sets makes evidence theory even closer to Shackle's intuitions. In fact, once belief functions are interpreted as measuring the degree of membership of propositions P to a fuzzy set Θ in terms of degrees of possibility, with $Bel(\mathrm{P}) = 0$ if P is impossible, $Bel(\mathrm{P}) = 1$ if P is perfectly possible and $Bel(\mathrm{P}) \in (0,1)$ if it is in between, evidence theory fits even Shackle's phraseology:

> We can go far in the analysis of decision by means of a mere possible-impossible dichotomy. Uncertainty as we ordinarily understand this word, however, includes a more subtle meaning of 'doubt' than the mere plurality of hypotheses placed in the 'possible' as against the 'impossible' class. There seems in the observed or reported working of our minds to be a faculty of adjudging degrees of possibility (...). (...) It is partly in the formation of such judgements that we think of the essentially new as being able to enter. It is here that there can be inspiration. This is the locus of creative decision.
>
> (Shackle, 1961, p. 12)

In the passage above Shackle understands beliefs as degrees of possibility, but he also understands creativity in terms of conceiving new possibilities, or "unpredictable hypotheses". Evidence theory understands creativity in decision-making as changing the frame of discernment, which amount to envisage new possibilities.

Obviously, evidence theory does not attempt to formalise the construction of a new frame of discernment. Shafer did single out a few rationales for expanding or contracting it, however:

> Like any creative act, the act of constructing a frame of discernment does not lend itself to thorough analysis. But we can pick out two considerations that influence it: (1) we want our evidence to interact in an interesting way, and (2) we do not want it to exhibit too much internal conflict.
> (...) Since it depends on what we are interested in, any judgment as to whether our frame is successful in making our evidence interact in an interesting way is a subjective one. But since interesting interactions can always be destroyed by loosening relevant assumptions and thus enlarging our frame, it is clear that our desire for interesting interaction will incline us towards abridging or tightening our frame.
> Our desire to avoid excessive internal conflict in our evidence will have precisely the opposite effect: it will incline us towards enlarging or loosening our frame. For internal conflict is itself a form of interaction - the most extreme form of it. And it too tends to increase as the frame is tightened, decrease as it is loosened.

(Shafer, 1976, p. 280)

Abridging, tightening, contracting a frame of discernment means to exclude some possibility from it, which amounts not to think to it anymore in order to focus on more interesting interactions. Likewise, enlarging, loosening or expanding a frame of discernment means to think new possibilities in order to adjust conflicting items of empirical evidence. For instance, by adding epicycles to the Ptolemaic theory medieval astronomers were able to adjust their frame of discernment to the results of empirical observations, while Newton's laws simplified and contracted the frame of discernment of celestial mechanics in order to derive interesting insights.

Shafer also warned that nothing warrants that an individual uses only one single frame of discernment at a time:

> Since we do find it necessary to adjust our assumptions to our circumstances, we use many different and incompatible frames of discernment in our practice of probable reasoning. In fact, we often consider many different frames more or less simultaneously. Even when our attention is sharply focused we tend to experiment to some extent at varying our assumptions, and as we shift the focus and level of generality of our attention we tend to vary these assumptions even more, sometimes going so far as to take for granted on one occasion something that might be the focus of our questioning on another. Consequently we will construct many different frames of discernment, which will vary in their success in organizing our experience and thus in their ultimate acceptance.

(Shafer, 1976, p. 281)

Possibly, studying the interactions between alternative frames is the key to understand how frames of discernment evolve with time. Dempster-Shafer rule can also be applied to the combination of alternative frames of discernment, though it suffers from assuming that evidence bodies are independent from one another (Ray and Krantz, 1996).

Lack of research on the evolution of frames of discernment is actually the weak point of evidence theory, a weakness that impaired the development of an alternative decision theory, able to overcome expected utility maximisation. Something has been done, however, and it is the subject of the next and last section.

5. Some Hints for an Alternative Decision Theory

Shackle developed a whole decision theory out of his concept of potential surprise. Its distinguishing feature is that only the worst and the best among the hypothetical outcomes of a decision are supposed to influence decision-making, contrary to expected utility maximisation where all possible consequences are weighed.

Shackle's decision theory is not a simple one, and its explanation would take an article in its own. J. Mars proposed a radical simplification of it, that was rejected by Shackle (Mars, 1950; Shackle, 1955, pp. 75-79). Unfortunately, because J.M. Ponsonnet recently took up Mars' suggestion again, demonstrating its equivalence to the original theory almost everywhere (Ponsonnet, 1996). It is Shackle's decision theory in the simplified Mars-Ponsonnet version that we shall expound and comment in the light of evidence theory.

Let each hypothesis that is entertained by a decision-maker at a certain point in time be endowed by some 'ascendancy' $\Phi(x, y)$, that depends both on its face value x (measured in terms of gains and losses) and on the potential surprise y the decision-maker attaches to it. However, if each hypothesis is characterised by a different potential surprise we can write $y = y(x)$, so that Φ can be written as a function of x only. Its typical shape is:

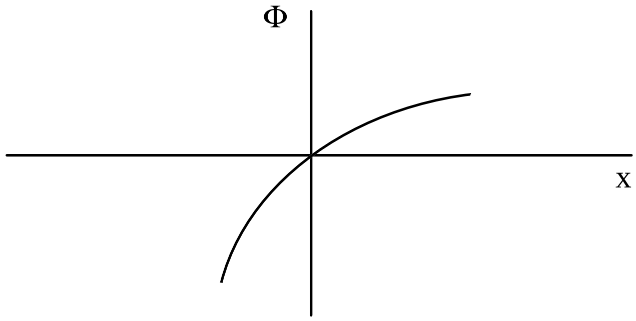

Fig. 1

Ascendancy is not just another name for utility. Shackle defines it as "attention arresting power, interestingness or ascendancy" (Shackle, 1961, p. 144), expressions that convey the idea of the decision-maker's mind focusing on some decision alternative; utility, on the contrary, simply measures how useful the outcome of an alternative is. Ascendancy entails the attractiveness of alternative courses of action in decision settings where each alternative will influence the future evolution to such a deep extent that that decision setting will never occur again and that alternative will never be available again.

Shackle assumed that decision-making ultimately consists of 1) focusing attention on the two hypothetical outcomes of each alternative that bear highest and lowest ascendancy, which he called the two "focus elements" of the ascendancy function, and 2) choosing among alternative courses of action according to preferences that are defined over their focus elements pairs. In its Mars-Ponsonnet's version, Shackle's decision theory boils down to the following two-steps algorithm:

1) For any action A, select the two hypotheses that maximise and respectively minimise ascendancy: $\overline{x}_A$: $\Phi(\overline{x}_A, y(\overline{x}_A)) = \max_{x_A} \Phi(x_A, y(x_A))$ and $\underline{x}_A$: $\Phi(\underline{x}_A, y(\underline{x}_A)) = \min_{x_A} \Phi(x_A, y(x_A))$. Hypotheses bearing outcomes $\overline{x}_A$ and $\underline{x}_A$ are also called 'focal gain' and 'focal loss', respectively.
2) Select a course of action by maximising $\Phi(A) = \Phi(\overline{x}_A) + \Phi(\underline{x}_A)$ with respect to A, where $\Phi(\overline{x}_A) \geq 0$ and $\Phi(\underline{x}_A) \leq 0$. The outcome of $\max_A \Phi(A)$ is the optimal course of action to pursue A^*.

How does this decision algorithm relate to Shafer's theory of evidence ?

Shafer himself never worried about decision-making, evidence theory being concerned with information representation only. However, an extension of von Neumann - Morgenstern expected utility to belief functions has been proposed by Jaffray (Jaffray, 1989). Belief function -based expected utility maximisation has interesting commonalties with Shackle's decision algorithm, as well as a fundamental difference that should not be overlooked.

Similarly to von Neumann and Morgenstern, who assumed the existence of a preference relation over probability distributions, Jaffray assumes the existence of a preference relation over belief functions. The utility associated to belief function *Bel* is:

$$U(Bel) = \sum_{A \in \Theta} m(A) v(A) \tag{7}$$

where utility function $v(A)$ is evaluated under the "sure thing" hypothesis $Bel(A) = 1$.

The trouble with (7) is that $v(A)$ depends on all outcomes $x \in A$; consequently, its calculation is prohibitive for a boundedly rational decision-maker. However, in an early formalisation attempt of Shackle's decision theory Arrow and Hurwicz established that, even assuming infinitely many unknown outcomes, there exists a function which depends only on the worst and the best among them, and which is able to yield the optimal course of action (Arrow and Hurwicz, 1972). Following a similar line of reasoning Jaffray shows that, under similar assumptions, there exists a function *w* such that (7) is equivalent to:

$$U(Bel) = \sum_{A \subset \Theta} m(A)\, w(\underline{x}_A, \overline{x}_A) \tag{8}$$

where $\underline{x}_A$ and $\overline{x}_A$ are the worst and the best possible outcomes in A. The similarity with Shackle's theory of decision-making is striking.

Nevertheless, utility *U* is a completely different object from ascendancy Φ. Just remark that in (8) we have the product of two magnitudes, while in Shackle's algorithm the ascendancy function alone is sufficient to evaluate alternative courses of action: any utility-based approach postulates that it is possible to disentangle the value of an outcome, expressed by its utility, from the belief that it may occur, be this belief expressed by means of probability theory or by means of evidence theory.

Shafer explicitly rejected this separation of value and belief that underlies any utility-based approach:

> My point is that the process of formulating and adopting goals creates a dependence of value and belief, simply because goals are more attractive when they are feasible.
>
> (Shafer, 1986a)

Shafer's claim reminds of that ancient Greek tale where a fox sees a bunch of grapes, tries in vain to reach it, and goes away saying that it's sour. Some argue that such a behaviour is much more widespread than it is commonly believed, and that it is based on an individual's desire to avoid cognitive dissonance between what he believes and what he actually does (Starmer, 1993): if it is too difficult to change one's behaviour (reaching the grapes), an individual is likely to change his beliefs (the grape is sour).

Shackle's "ascendancy", this "attention arresting power" that is defined independently of preferences but includes human wishes, seems to follow this tradition. Shafer's evidence theory does not provide any formalisation of Shackle's "ascendancy", if ascendancy is understood in this way. It is a young theory, however, which a few mathematicians are slowly bringing to formal perfection and whose diffusion among economists this essay would like to start.

Notes

(1) Mental habits stemming from probability theory led Shafer to call the numbers m_j "basic probability assignments". This name is misleading: although numbers m_j are required to sum up to unity, they are not probabilities because they are defined on a set - the frame of discernment - which is not an algebra.

(2) The coin-tossing setting has been superimposed to Shafer's example in order to ease understanding. Moreover, Shafer's original example has been simplified in that any reference to unequal probabilities (unfair coins) has been expunged.

(3) For the sake of brevity, in this exposition many interesting issues contained in the original example have been expunged.

(4) Shafer actually used the symbol "A" where I wrote "P". I did so in order to avoid confusion with the empirical evidence bearing on $A_1, A_1, \ldots$.

References

Arrow, K.J. and Hurwicz, L. 1972. An Optimality Criterion for Decision-Making under Ignorance, in Carter, C.F. and Ford, J.L. (eds.), *Uncertainty and Expectations in Economics*, Clifton, Augustus M. Kelley Publishers.

Choquet, G. 1953-54. Theory of Capacities, *Annales de l'Institut Fourier*, vol. 5, 131-295.

De Finetti, B. 1931. Sul Significato Soggettivo della Probabilità, *Fundamenta Mathematicae*, vol. 17, 298-329.

De Finetti, B. 1970. *Teoria della Probabilità*, Torino, Einaudi.

Dempster, A.P. 1966. New Methods for Reasoning towards Posterior Distributions based on Sample Data, *The Annals of Mathematical Statistics*, vol. 37, 355-74.

Dempster, A.P. 1967. Upper and Lower Probabilities Induced by a Multivalued Mapping, *The Annals of Mathematical Statistics*, vol. 38, 325-38.

Dempster, A.P. 1968. A Generalization of Bayesian Inference, *Journal of the Royal Statistical Society, Series B (methodological)*, vol. 30, 205-47.

Dubois, D. and Prade, H. 1988. *Possibility Theory*, New York, Plenum Press.

Ellsberg, D. 1961, Risk, Ambiguity, and the Savage Axioms, *The Quarterly Journal of Economics*, vol. 75, 643-69.

Ford, J.L. 1987. *Economic Choice under Uncertainty*, Aldershot, Edward Elgar.

Ford, J.L. and Ghose, S. 1994a. 'Belief, Credibility, Disbelief, Possibility, Potential Surprise and Probability as Measures of Uncertainty: Some Experimental Evidence', Discussion Paper No. 94-22, University of Birmingham, Department of Economics.

Ford, J.L. and Ghose, S. 1994b. 'Shackle's Axioms for Potential Surprise and Belief: A Preliminary Experimental Assessment', Discussion Paper No. 94-21, University of Birmingham, Department of Economics.

Ford, J.L. - Ghose, S. 1995a. 'Ellsberg's Urns, Ambiguity and Measures of Uncertainty: Some Experimental Evidence', Discussion Paper No. 95-12, University of Birmingham, Department of Economics.

Ford, J.L. and Ghose, S. 1995b. 'Shackle's Theory of Decision-Making Under Uncertainty: The Findings of a Laboratory Experiments' Discussion Paper No. 95-08, University of Birmingham, Department of Economics.

Ford, J.L. and Ghose, S. 1995c. 'The Primitive Uncertainty Construct: Possibility, Potential Surprise, Probability and Belief: Some Experimental Evidence', Discussion Paper No. 95-07, University of Birmingham, Department of Economics.

Gilboa, I. 1987. Expected Utility with Purely Subjective Non-Additive Probabilities, *Journal of Mathematical Economics*, vol. 16, 65-88.

Gilboa, I. and Schmeidler, D. 1993. Updating Ambiguous Beliefs, *Journal of Economic Theory*, vol. 59, 33-49.

Good, , I.J. 1952. Rational Decisions, *Journal of the Royal Statistical Society, Series B (methodological)*, vol. 14, 107-14.

Good, I.J. 1962. Subjective Probability as the Measure of a Non-Measurable Set, in Nagel, E., Suppes, P. and Tarski, A. (eds.), *Logic, Methodology and Philosophy of Science*, Stanford, Stanford University Press.

Harmanec, D. and Klir, G.J. 1994. On Modal Logic Interpretation of Dempster-Shafer Theory of Evidence, *International Journal of Intelligent Systems*, vol. 9, 941-51.

Hey, J.D. 1985. The Possibility of Possibility, *Journal of Economic Studies*, vol. 12, 70-88.

Jaffray, J.Y. 1989. Linear Utility Theory for Belief Functions, *Operations Research Letters*, vol. 8, 107-12.

Jaffray, J.Y. 1992. Bayesian Updating and Belief Functions, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, 1144-52.

Jaffray, J.Y. and Philippe, F. 1997. On the Existence of Subjective Upper and Lower Probabilities, *Mathematics of Operations Research*, vol. 22, 165-185.

Kahneman, D. and Tversky, A. 1979. Prospect Theory: An Analysis of Decision Under Risk, *Econometrica*, vol. 47, 263-91.

Katzner, D.W. 1986. Potential Surprise, Potential Confirmation, and Probability, *Journal of Post-Keynesian Economics*, vol. 9, 58-78.

Klir, G.J. and Harmanec, D. 1994. On Modal Logic Interpretation of Possibility Theory, *International Journal of Uncertainty, Fuzzyness and Knowledge-Based Systems*, vol. 2, 237-245.

Klir, G.J. and Yuan, B. 1995. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Upper Saddle River, Prentice Hall.
Koopman, B.O. 1940a. The Axioms and Algebra of Intuitive Probability, *Annals of Mathematics*, vol. 41, 269-92.
Koopman, B.O. 1940b. The Bases of Probability, *Bulletin of the American Mathematical Society*, vol. 46, 763-74.
Koopman, B.O. 1941. Intuitive Probabilities and Sequences, *Annals of Mathematics*, vol. 42, 169-87.
Kries, J. von 1886. *Die Principien der Wahrscheinlichkeitsrechnung*, Tübingen, J.C.B. Mohr (Paul Siebeck).
Mars, J. 1950. A Study in Expectations. Reflections on Shackle's "Expectation in Economics", *Yorkshire Bulletin of Economic and Social Research*, vol. **2**(2), 63-98 and **2**(3), 1-35.
Ponsonnet, J.M. 1996. The Best and the Worst in G.L.S. Shackle's Decision Theory, in Schmidt, C. (ed.), *Uncertainty in Economic Thought*, Cheltenham, Edward Elgar.
Ramsey, F.P. 1931. Truth and Probability, in Braithwaite, R.B. (ed.), *The Foundations of Mathematics and Other Logical Essays*, London, Routledge & Kegan Paul.
Ray, B.K. and Krantz, D.H. 1996. Foundations of the Theory of Evidence: Resolving Conflict Among Schemata, *Theory and Decision*, vol. 40, 215-34.
Resconi, G., Klir, G.J., Harmanec, D. and St.Clair, U. 1996. Interpretations of Various Uncertainty Theories Using Models of Modal Logic: A Summary, *Fuzzy Sets and Systems*, vol. 80, 7-14.
Resconi, G., Klir, G.J. and Pessa, E. 1999. Conceptual Foundations of Quantum Mechanics: The Role of Evidence Theory, Quantum Sets and Modal Logic, *International Journal of Modern Physics, series C*, vol. 10, 29-62.
Resconi, G., Klir, G.J. and St.Clair, U. 1992. Hierarchical Uncertainty Metatheory Based Upon Modal Logic, *International Journal of General Systems*, vol. 21, 23-50.
Resconi, G., Klir, G.J., St.Clair, U. and Harmanec, D. 1993. On the Integration of Uncertainty Theories, *International Journal of Uncertainty, Fuzzyness and Knowledge-Based Systems*, vol. 1, 1-18.
Sarin, R. and Wakker, P. 1992. A Simple Axiomatization of Nonadditive Expected Utility, *Econometrica*, vol. 60, 1255-72.
Sarin, R. and Wakker, P. 1994. Gains and Losses in Nonadditive Expected Utility, in Munier, B. and Machina, M.J., *Models and Experiments in Risk and Uncertainty*, Dordrecht, Kluwer Academic Publishers.
Savage, L.J. 1954. *The Foundations of Statistics*, New York, Wiley.
Schmeidler, D. 1989. Subjective Probability and Expected Utility without Additivity, *Econometrica*, vol. 57, 571-87.
Shackle, G.L.S. 1955. *Uncertainty in Economics and Other Reflections*, Cambridge, Cambridge University Press.
Shackle, G.L.S. 1961. *Decision, Order and Time in Human Affairs*, Cambridge, Cambridge University Press.
Shafer, G. 1976. *A Mathematical Theory of Evidence*, Princeton, Princeton University Press.
Shafer, G. 1979. Allocations of Probability, *The Annals of Probability*, vol. 7, 827-839.
Shafer, G. 1981. Constructive Probability, *Synthese*, vol. 48, 1-60.
Shafer, G. 1982. Belief Functions and Parametric Models, *Journal of the Royal Statistical Society, Series B (methodological)*, vol. 44, 322-352.
Shafer, G. 1986a. Savage Revisited, *Statistical Science*, vol. 1, 463-501.
Shafer, G. 1986b. The Combination of Evidence, *International Journal of Intelligent Systems*, vol. 1, 155-79.
Shafer, G. 1990. Perspectives on the Theory and Practice of Belief Functions, *International Journal of Approximate Reasoning*, vol. 4, 323-62.
Shafer, G. 1992. Rejoinders to Comments on "Perspectives on the Theory and Practice of Belief Functions", *International Journal of Approximate Reasoning*, vol. 6, 445-80.
Shafer, G. and Tversky, A. 1985. Languages and Designs for Probability Judgment, *Cognitive Science*, vol. 9, 309-39.
Starmer, C. 1993. The Psychology of Uncertainty in Economic Theory: A Critical Appraisal and a Fresh Approach, *Review of Political Economy*, vol. 5, 181-96.
Tversky, A. and Kahneman, D. 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty*, vol. 5, 297-323.
Walley, P. 1991. *Statistical Reasoning with Imprecise Probabilities*, London, Chapman and Hall.